一种基于信用的改进 PBFT 高效共识机制

徐治理 [1,2]，封化民 [1,2]，刘　飚 [1]

(1. 北京电子科技学院，北京 100070; 2. 西安电子科技大学 通信工程学院，西安 710071)

摘　要：基于联盟链的特点和要求，分析共识机制的协议性能和资源的使用，在 PBFT（practical byzantine fault tolerance，实用拜占庭容错技术）的基础上，对共识过程中的协议进行了针对性的改进。新共识机制中，首先引入了信用评价，配合简化的一致性协议促进系统进入良性循环；修改检查点协议，使节点能动态的加入、离开系统，提高系统灵活性。实验结果表明，新的共识机制能在长期运行中缩短交易确认时间、减少通信资源的使用、提高系统的效率。

关键词：联盟链；共识机制；实用拜占庭容错算法

中图分类号：TP301　　doi: 10.3969/j.issn.1001-3695.2018.03.0195

Study of highly efficient PBFT consensus mechanism based on credit

Xu Zhili[1, 2], Feng Huamin[1, 2], Liu Biao[1]

(1. *Beijing Electronic Science & Technology Institute*, *University*, *Beijing* 100070, *China*; 2. *School of Telecommuncations Engneegring*, *Xidian University*, *Xi'an* 710071, *China*)

Abstract: Based on the characteristics and requirements of the consortium blockchain, this paper analyzed the protocol performance of the consensus mechanism and the spending of resources. Based on PBFT (practical byzantine fault tolerance) , this paper targeted the agreement in the consensus process for improvement. In the new consensus mechanism, it first introduced the credit evaluation, and used the simplified coherence protocol to promote the system to enter into a virtuous circle. It modified the checkpoint protocol that the nodes could dynamically join and leaved the system and improve the system flexibility. The experimental results show that the new consensus mechanism can shorten transaction confirmation time, reduce the use of communication resources, and improve system efficiency in the long running.

Key words: consortium blockchain; consensus mechanism; PBFT

0　引言

区块链技术最早作为比特币错误!未找到引用源。的支撑技术被提出。比特币作为一种数字加密货币，与传统货币相比，比特币将自己定位成为一种信用去中心化、不受权威机构控制的系统，区块链技术为分布式系统提供参与者之间的信任。区块链采用了密码学、计算机和通信领域的各项技术，使用非对称加密、时间戳、共识机制和点对点通信，解决现有的中心化信用机构的效率低、成本高和数据所有权被垄断的问题。区块链技术被认为是移动互联网之后新的信息技术发展方向，将促进信用社会的建立，促使目前的信息互联网向价值互联网转变[1]。

区块链最早的应用是以比特币为代表的数字货币，这方面的应用也被了解得最为广泛，但它们功能局限于货币交换，仅体现了区块链去中心化的特性。为了实现更智能的应用，业界推出了可编程区块链系统，比如以太坊（etherum）可由各节点运行被称为“智能合约”的图灵完备脚本语言 EtherScript 错误!未找到引用源。，实现自动控制管理链上资产的功能。区块链系统的发展方向将会超越金融范围，实现智能物联网、能源互联网等数据服务平台。

最早被提出的区块链参与形式被称为公有链，系统的结构完全去中心化，任何用户都能参与并获取链上存储的所有数据错误!未找到引用源。，但由于系统开放性的限制，用户的隐私和监管都存在不足，尤其是金融领域的相关应用难以满足效率需求。因此，以 R3、Hyperledger 为代表的行业内部联盟链成为区块链和行业结合的发展方向错误!未找到引用源。。参与联盟链的节点是可验证身份的，并且在联盟内部可以实现平等、可信、高效的操作，如金融领域不同机构间的汇兑和清算。

本文针对联盟链的应用场景，在 PBFT 共识机制错误!未找到引用源。的基础上进行改进，参考 PoW（proof of work，工作量证明）的思想，将参与者行为作为工作量证明充分利用，在系统的长期运行中维持良性循环，减少通信资源的占用提高系统效率。

收稿日期：2018-03-13；**修回日期：**2018-04-23

作者简介：徐治理（1991-），男，河北沧州人，硕士研究生，主要研究方向为区块链（xuzlfight@qq.com）；封化民（1963-），男，教授，博士，主要研究方向为区块链、信息安全、机器学习；刘飚（1981-），男，博士，讲师，主要研究方向为侧信道攻击、机器学习.

1　共识机制介绍

去中心化的区块链系统如何高效的达成分布式系统的一致性是决定性能好坏的重要重要问题错误!未找到引用源。系统状态的决定权分散在每一个成员手中，成员规模越大，达成共识的计算和通信成本越高。分布式系统中如果仅存在被动错误，如数据包丢失和延迟，可以通过 Raft 和 Paxos 等算法解决一致性问题错误!未找到引用源。；如果分布式系统中节点之间属于互相不了解的参与者，受到利益的驱使，则还可能产生拜占庭错误，即存在节点主动向其他节点发送错误信息。在这种系统中需要使用有拜占庭容错能力的共识算法[8]。

1.1　工作量证明

PoW 共识机制中，各个节点同时计算一个求解困难且验证容易的密码学问题，最先解出问题的节点获得生成当前区块的权利，并获得一定的收益错误!未找到引用源。收益可以鼓励记账者积极竞争获取记账机会，不同节点间的博弈又促使了记账者记录正确的信息，只要诚实的参与者占系统中的大多数，区块链中的记录就可以被信任，很好地防止了“女巫攻击”错误!未找到引用源。PoW 共识在比特币的提出和发展中有重要的意义，在设计之初很好地保障了比特币的安全性和公平性。随着区块链的发展 PoW 暴露出很多问题：在能耗方面存在巨大的浪费，大量电力被用来计算无意义的问题；出块间隔太长，导致交易确认效率低；专用芯片的出现使少数机具有构垄断区块生成的能力，记账结果的正确性和整个系统的抗攻击能力都受到影响。

1.2　权益证明

PoS（proof of stake，权益证明）共识机制是为了解决 PoW 的资源浪费提出的替代性方案错误!未找到引用源。PoS 引入了一个新的概念“币龄”（coin days），币龄是用户持有的系统代币的数量和持有时间的乘积，系统根据矿工持有的币龄降低挖矿难度，矿工生成区块后会消耗相应的币龄，从而使用权益证明代替工作量。但 PoS 也存在部分缺点：动态的挖矿难度会影响出块的平稳程度；币龄的积累可以离线完成，无法促进节点参与共识过程；鼓励用户囤积代币，减弱了系统代币的流动性。

1.3　授权股份证明

DPoS（delegated proof-of-stake）对中心化进行了适当妥协，使用了代议制的结构错误!未找到引用源。，所有参与共识的节点投票选出少量见证节点完成区块生成，因为直接参与共识的节点减少，所以达成共识的速度大大增加。当见证节点不能正确的记录交易信息、及时同步新的区块时，其他用户可以投票将其替换，生成区块带来的收益将促使被选中的见证节点认真履行共识过程中的职责，维持自己的权限。

1.4　实用拜占庭容错技术

PBFT 是解决存在拜占庭错误节点的分布式系统一致性问题的通用方案。区块链可以视为一个异步的分布式系统，其数据结构适用于状态机拜占庭容错的使用场景错误!未找到引用源。PBFT 共识主要由一致性协议、视图更换协议和检查点协议组成。区块生成过程中，节点运行一致性协议，通过点对点通信获取其他状态后决定自身状态，通过类似投票的过程完成对每个节点区块链数据的统一；在一致性协议运行出现超时故障时，系统通过视图更换协议替换故障节点，保证系统运行稳定；检查点协议定期清除过期交互数据减轻节点存储负担，并负责定期检查系统状态是否统一，对不一致的节点进行同步。区块链系统 Hyperledger Fabric 已经使用了 PBFT 算法实现共识。PBFT 节点间存在大量的点对点通信，协商达成状态一致，对于通信资源的占用较大，因此优化通信过程提高共识效率的研究具有重要意义。

2　改进方案核心协议

本文基于联盟链的应用场景，对原有的 PBFT 进行了部分改进，引入行为积分和分级机制，改进的共识机制称为 CPBFT（credit practical byzantine fault tolerance，信用拜占庭协议）。其中将节点行为作为工作量证明，与优化后的通信过程相结合，通过实现系统的良性循环提高效率减少通信成本。联盟链要求参与者经过身份认证才能加入，可信程度和稳定性都比公有链更有保证，参与者的规模也相对较小，多项式的通信复杂度在实际中也能满足。

PBFT 共识主要由一致性协议、视图切换协议和检查点协议组成。在 CPBFT 共识机制中，视图切换协议增加了出块失败信用惩罚；检查点协议中增加了动态成员和信用排序内容。

2.1　PBFT 共识定义

使用 PBFT 共识机制的区块链系统拥有 n 台服务器，系统节点包括正常节点和存在拜占庭错误的节点，系统能承受最大拜占庭节点的数目为 f，系统正确运行要求 $n \geq 3f+1$，正常节点数目至少是拜占庭节点数目的两倍。PBFT 为保证系统的一致性和安全性达成共识过程中要求：

a）正常节点存储的区块链数据和交易信息正确且一致；

b）如果确认区块数据正确，正常节点必须接受该区块。

节点间为了达成共识需要广播和交换消息，这些消息称为证书。证书中包括该证书编号、发送者身份信息、协商内容所属协议等内容。节点收到信息后会检查信息并记录在日志中。区块链中区块的数量被称为区块高度，区块的高度也可以指某个区块在链条中的位置。

2.2　一致性协议和简化一致性协议

区块链系统中每隔一段时间，一定数量的交易或者账户状态变化被打包记录进区块中。系统中节点通过一致性协议保证记录的区块信息正确并相同。协议中存在主节点（primary）和从节点（replica）两种角色。一次共识中主节点只存在一个，负责对一段时间内接收到的交易进行验证，通过验证的交易将被打包进区块。参与共识的节点会有从 0 到 $R-1$ 的互不相同的编号，主节点选择公式为 $p=(h+v)\bmod(|R|)$。其中：p 是被选中节点编号；v 是从零开始逐渐递增的视图编号；$|R|$ 是参与

共识的节点总数。

一次完整的区块生成共识发起并完成需要三个阶段的通信过程如图 1 所示。具体步骤如下：

a）主节点生成 Pre-Prepare 证书，其中包括新区快、证书时间戳和主节点签名等。主节点将 Pre-Prepare 证书发送给从节点，之后主进入 Prepare 状态。

b）从节点收到 Pre-Prepare 证书后，如果是第一次收到该证书从节点进入 Prepare 状态，并转发该证书给其他从节点。

c）节点收到其他节点发来的证书，会校验证书内信息，包括区块内交易的正确性、区块头信息正确性和区块高度等。如果认可该区块，向发送来证书的节点回复认可反馈。一个节点收到包括自己的 $2f+1$ 个认可反馈，表明该区块被加入区块链尾端，这条证书进入 Commit 状态。

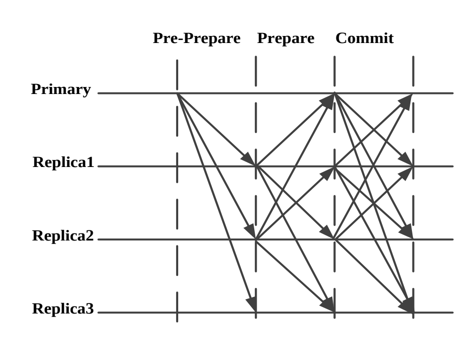


图 1　一致性协议完整交互过程

图 1 中 primary 为主节点；replica1 和 replica2 为诚实且无延迟的从节点，replica3 存在被动错误，即使从节点 Replica3 存在拜占庭错误，根据相同的投票过程协议仍能正确执行。

完整的一致性协议需要完成两次复杂度为 $O(n^2)$ 的通信过程，复杂度较高。所以，为了简化通信过程，本文参考混合群组拜占庭容错错误!未找到引用源。对没有拜占庭错误的情况下进行优化。简化过程如图 2 所示。具体内容为：

a）主节点发送给所有从节点 Pre-Prepare 证书，从节点收到后如果认可证书内容，回复认可信息；

b）主节点如果收到 $3f$ 个认可信息，将认可信息打包再发送给所有从节点，从节点可以验证其他节点的认可信息是否正确。如果正确证明所有节点都接收该区块信息，所有节点进入 commit 状态，将新区块加入区块链尾端。

c）如果主节点没有收到所有的认可信息，则进入完整的一致性协议流程。

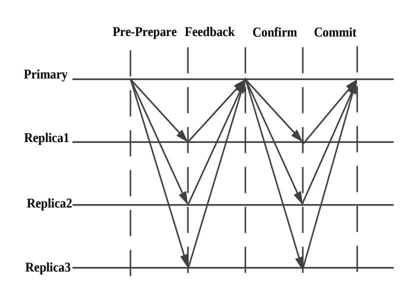


图 2　一致性协议简化交互过程

成功完成区块生成后，将打包进区块中的交易从待确认列表中移除，视图编号增加 1，进入到下一个区块生成过程中，循环一定的周期数后会进入检查点协议更新节点信用值和节点顺序。

2.3　视图切换协议

视图是 PBFT 共识机制中节点关系的定义，视图的编号记为 v，视图中节点有不同的编号，每个视图有一个主节点。一致性协议中主节点拥有记录交易到区块中的核心权利，主节点产生错误会导致出块停止，视图切换协议在主节点故障时完成变更主节点的任务，保证维持系统的运行，视图变更后视图编号增加 1。

视图切换协议由从节点触发，以区块链中最新区块的时间戳 T 为起始时间，配合一致性协议有两个超时触发条件：在有限的时间 ΔT_1 内没有收到新的主节点 Pre-Prepare 协议；在有限时间 ΔT_2 内没有完成新的区块生成。其中 $\Delta T_1 < \Delta T_2$。满足上述两个条件的任意一个则可以认为主节点故障，此时需要进行视图切换。

视图切换需要节点间的交互通信，执行过程如下：

a）从节点开始执行视图切换协议后进入视图 $v+1$，发送 View-Change 证书给所有节点，其中包括最新区块编号和摘要信息，和新的视图编号及主节点。

b）节点如果包括自己收到 $2f+1$ 条 View-Change 证书，则发送给视图 $v+1$ 中主节点 View-Change-Ack 证书。同时清除 T 之后收到的一致性协议证书，等待新的主节点发起一致性协议。

c）主节点根据自己存储的区块链数据开始新一轮一致性协议。

视图切换过程会造成交易确认的停止，所以需要尽量避免主节点错误，通过信用累积，多次顺利完成区块生成的节点将会有更多的机会担任主节点，提高系统效率。

2.4　检查点协议和信用分级协议

理想情况下所有节点能及时地完成系统中的各项交互任务保持一致性，但实际中部分节点可能由于本身故障或网络问题落后于其他节点，需要一个周期性协议同步整个系统来防止节点不一致累积导致系统故障。检查点协议在检查一致性状态后，对已经确认的区块相关的证书进行清除，减少节点存储压力。本文还在检查点协议中扩展了动态增删节点和信用分级排序功能。

检查点协议执行时间间隔记为 CT，每次执行节点向其他节点索要区块链状态信息，一旦发现与大多数节点不一致，则主动向其他节点索要从上一个检查点开始的区块信息。同步完成后，节点更新本地交易列表清除已经被记录的交易，清除在最新区块时间戳之前的证书。

系统中根据节点行为产生信用分数，根据信用分数的评级

分 A、B、C 三等，三类节点能完成的协议如表 1 所示。

表 1　节点权限比较

信用级别	权限		
	担任主节点	担任主节点	参与检查点协议
A	√	√	√
B		√	√
C			√

刚加入系统的节点为 C 类节点，在同步完区块后成为 B 类节点。节点在系统中参与完成一次区块生成共识加 1 分；成为主节点成功完成区块生成会收取一定的实际收益，但该轮共识不加信用分；未能成功生成区块，扣 5 分。节点会因为自身行为在 A 和 B 两类节点中变化。根据系统中节点的多少，以 CT 的整数倍时间周期执行信用分级协议，完成信用信息的更新，A 类节点获得视图编号用来参与主节点选择。信用分级协议也需要一次三阶段的通信共识，但在长期统计来看，配合简化一致性协议能大大减少通信开销，提高系统效率。

3　性能分析与实验

PoW、PoS、DPoS 等都是为了公有链应用环境而设计的，在应用到联盟链中时存在不适用的问题，联盟链的共识机制需要针对联盟链用户较少但对效率要求更高的特点进行选择和改进。PoW 和 PoS 基于证明的设计是为了防止攻击者恶意注册多个虚假身份提高自己在共识中的份额。但联盟链中节点的加入需要进行申请和身份认证，不需要再进行额外的证明，尤其工作量证明还要消耗额外的能源，所以联盟链不需要共识机制中的证明过程。DPoS 虽然与 PBFT 同样基于投票达成共识，但选举见证节点代表用户的结构更适合规模较大的系统，联盟链中用户规模较小，采用 PBFT 能更有效地对拜占庭节点的行为进行反馈，提高共识效率，也促进了系统的去中心化。本文提出的 CPBFT 在此基础上，通过对节点行为的记录和打分，通过限制错误节点权限进一步提高了共识效率。但由于 CPBFT 与 PBFT 一样达成一致需要大量的通信，所以仅适用于规模较小的联盟链系统。

本文对比 PBFT 共识机制和本文提出的 CPBFT 共识机制的运行效率，以交易确认速率为评价指标，测试在拜占庭错误节点占总节点中不同比例和不同运行时间下两种算法的性能差异。采用 PBFT 共识机制的区块链系统中，交易被打包进区块就可以被认为得到了确认。交易确认速率为单位时间能打包进区块的交易数量平均值，如式（1）所示。

$$TPS = transcations/\Delta t \quad (1)$$

其中：$transcations$ 为一段时间内包含进区块链的交易数；Δt 为记录时间，一般为区块生成时间的整数倍。

3.1　容错性能比较

系统中拜占庭错误节点所占比例是对系统性能影响最大的因素，CPBFT 并没有引入更严格的情景假设，所提供的容错能力与 PBFT 相同，均为 $f = \lfloor (n-1)/3 \rfloor$。图 3 分别为使用 PBFT 和 CPBFT 共识机制分别运行 0～10 min 钟平均每分钟交易确认速率情况。系统在总共设置 301 个节点。最大 100 个错误节点环境中，系统中错误节点随机变化但不会超过最大值。对比图 3 可以看到，在相同的系统环境中，短时间内 CPBFT 和 PBFT 能达到相同的效率。

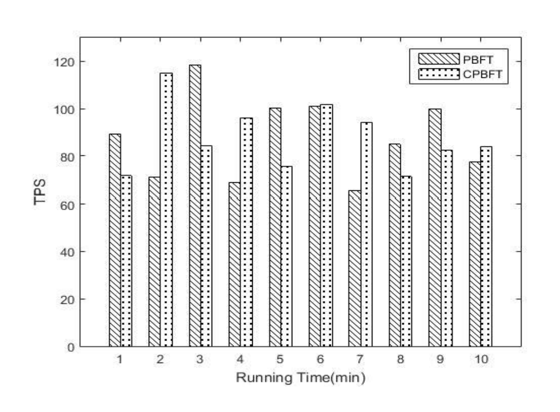


图 3　PBFT 与 CPBFT 交易吞吐量对比

3.2　运行效率比较

CPBFT 设计的目的之一就是提高系统长期运行效率。随着系统运行高错误率节点信用打分降低，低主节点错误率和简化一致性协议及激励相配合，在错误节点上限范围内，能比 PBFT 更高效地完成区块生成。图 4 为一段较长时间的 PBFT 和 CPBFT 的交易确认速率变化。系统采用图 4 中同样的设置。可以看到主节点错误率下降后系统的交易吞吐量有明显增加。

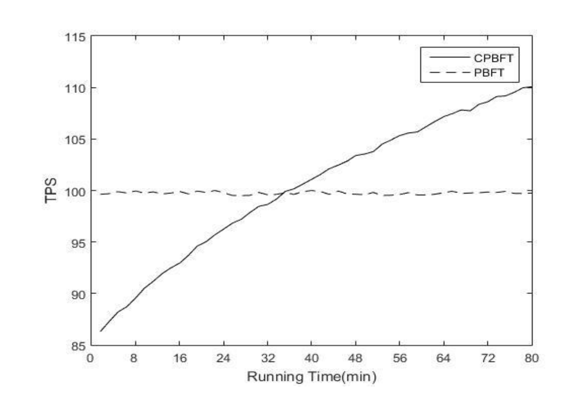


图 4　PBFT 与 CPBFT 两种共识机制长期运行吞吐率比较

3.3　通信开销验证

PBFT 及其衍生的共识机制存在的问题就是共识过程需要大量的节点间通信，CPBFT 中检查点协议由于加入了信用评分，需要额外的通信过程对信用评分达成一致，系统开始运行的短时间内或者系统中错误节点较少时，会增加传输数据量，但在错误节点较多的场景中，能减少视图切换协议的调用，在另一方面降低了数据量。图 5 为在图 4 的系统设置下，节点间数据传输量的比较。图中横轴为系统持续运行时间，纵轴为生成一个区块需要的复杂度为 $O\left(n^2\right)$ 的点对点通信次数。PBFT 没有运行中的优化机制，生成区块平均通信量没有变化；

CPBFT 随着主节点错误率下降，通信效率逐渐提高。

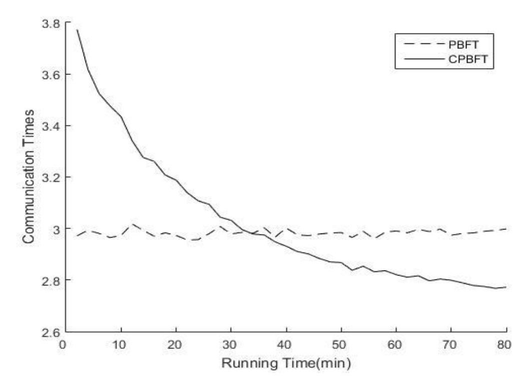


图 5　PBFT 与 CPBFT 生成单位区块通信量对比

4　结束语

基于投票的 PBFT 共识机制很好地解决了能源消耗和避免趋向中心化的问题，但节点间通信的网络开销是需要解决的问题。本文通过提出信用机制配合 PBFT 中的检查点协议，提出了一种高效动态的 CPBFT 共识机制，减少共识达成过程中的多次交互，提供了一种降低网络开销的解决方案，最后通过 MATLAB 进行仿真对系统的性能进行对比分析。

区块链作为创造信任的机器，是未来信息互联网向价值互联网发展中的关键技术。共识机制目前已经成为区块链系统性能的瓶颈。共识机制的目的是在互不信任的分布式系统节点间高效证明行为正确性和达成系统一致性。工作量证明、权益证明和基于投票的共识机制在不同的使用范围都提现了各自的优势。如何结合区块链系统应用场景，融合不同的共识机制的优势，是设计共识机制中值得关注的问题。

参考文献：

[1] Natamoto S. Bitcoin: a peer-to-peer electronic cash system [EB/OL]. (2009) [2018-03-01]. https://bitcoin. org/bitcoin. pdf.

[2] 邵奇峰，金澈清，张召，等. 区块链技术：架构及进展 [J]. 计算机学报，2018, 41 (5): 969-988. (Shao Qifeng, Jin Cheqing, Zhang Zhao, *et al.* Blockchain: architecture and research progress [J]. Chinese Journal of Computers, 2018, 41 (5): 969-988.)

[3] Buterin V. A next-generation smart contract and decentralized application platform [EB/OL]. (2014) [2018-03-01]. https://github. com/ethereum/wiki/wiki/White-Paper.

[4] 常文军，孙超平，胡鑫. 基于分布式偏好关系的多属性决策方法及应用 [J]. 计算机应用研究，2017, 34 (12): 3693-3697, 3707. (Chang Wenjun, Sun Chaoping, Hu Xin. Multiple attribute decision making method based on distributed preference relations and its application [J]. Application Research of Computers, 2017, 34 (12): 3693-3697, 3707.)

[5] Hull R. Blockchain: distributed event-based processing in a data-centric world: extended abstract [C]// Proc of ACM International Conference on Distributed and Event-Based Systems. 2017: 2-4.

[6] Castro M. Practical byzantine fault tolerance and proactive recovery [C]// Proc of Symposium on Operating Systems Design and Implementation. [S. l.] : USENIX Association, 1999: 173-186.

[7] Lamport L. Paxos made simple [J]. ACM Sigact News, 2016, 32 (4): 18-25.

[8] 张伯阳，张晓，李阿妮，等. 云存储系统可扩展性评测研究 [J]. 计算机应用研究，2017, 34 (7): 1957-1961, 1965. (Zhang Boyang, Zhang Xiao, Li Ani, *et al.* Research on scalability evaluation in cloud storage system [J]. Application Research of Computers, 2017, 34 (7): 1957-1961, 1965.)

[9] 邹均，张海宁，唐屹，等. 区块链技术指南 [M]. 北京：机械工业出版社，2016: 109-112. (Zou Jun, Zhang Haining, Tang Yi, *et al.* Blockchain technology guide [M]. Beijing: Chine Machine Press, 2016: 109-112.)

[10] Douceur J R. The sybil attack [C]// Proc of International Workshop on Peer-to-Peer Systems. Berlin: Springer, 2002: 251-260.

[11] 长铁，韩锋，杨涛，等. 区块链：从数字货币到信用社会 [M]. 北京：中信出版集团股份有限公司，2016: 27-29. (Chang Jia, Han Feng, Yang Tao, *et al.* Blockchain: from digital currency to credit society [M]. Beijing: CITIC Publishing Group, 2016: 27-29.)

[12] Houy N. It will cost you nothing to'kill'a proof-of-stake crypto-currency [J]. Social Science Electronic Publishing, 2014, 34 (2): 1038-1044.

[13] 范捷，易乐天，舒继武. 拜占庭系统技术研究综述 [J]. 软件学报，2013, 24 (6): 1346-1360. (Fan Jie, Yi Letian, Shu Jiwu. Research on the technologles buzantine system [J]. Journal of Software, 2013, 24 (6): 1346-1360.)

[14] 袁勇，王飞跃. 区块链技术发展现状与展望 [J]. 自动化学报，2016, 42 (4): 481-494. (Yuan Yong, Wang Feiyue. Blockchain: the state of the art and future trends [J]. Acta Automatica Sinica, 2016, 42 (4): 481-494.)

[15] Cowling J, Myers D, Liskov B, *et al.* HQ replication: a hybrid quorum protocol for byzantine fault tolerance [C]// Proc of Usenix Symposium on Operating Systems Design and Implementation. USENIX Association. 2006: 13-13.